

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 21, 2005, 12:37:33 ; Search time 169 Seconds

(without alignments)  
769,494 Million cell updates/sec

Title: US-10-809-816A-2

Perfect score: 1799  
Sequence: 1 MKQLCVLLVCSSAVAGLHK.....MARNKGNHGIGAFSPYPERI 331

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A\_Geneseq\_21: \*  
1: geneseqp19808:\*  
2: geneseqp19908:\*  
3: geneseqp20008:\*  
4: geneseqp20018:\*  
5: geneseqp20028:\*  
6: geneseqp20038:\*  
7: geneseqp20038:\*  
8: geneseqp20048:\*  
9: geneseqp20058:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1747	97.1	331	7	ADP65168 Human cat
2	1747	97.1	331	9	ADY14944 PRO polyp
3	1747	97.1	331	9	ADY14946 PRO polyp
4	1747	97.1	331	9	AEC01587 Human cat
5	1744	96.9	331	7	ADDA5416 Human cat
6	1744	96.9	331	7	ADDS6902 Human pro
7	1744	96.9	331	8	ADJ71695 Human NOV
8	1741	96.8	331	7	ADDS5927 Human cat
9	1740	96.7	331	3	AAVS9634 Human cat
10	1738	96.6	331	9	ADY19784 PRO polyp
11	1734	96.4	331	7	ADDS5932 Human cat
12	1731	96.2	331	7	ADDS5934 Human cat
13	1724	95.8	331	7	ADDS5933 Human cat
14	1701.5	94.6	330	6	AAO30451 Human cat
15	1631	90.7	313	8	ABM83324 Human dia
16	1631	90.7	313	8	ABM83323 Human dia
17	1600	88.9	331	6	ABR56160 Canine ca
18	1580	87.8	331	6	ABG73437 Human cat
19	1388	77.2	330	7	ADDA69537 Human hea
20	1309.5	72.8	330	7	ADDA5414 Rat Prote
21	1309.5	72.8	330	7	ADDS6900 Rat Prote
22	1145	63.6	217	7	ADDS5931 Human mat
23	1144	63.6	222	8	ADJ71697 Human NOV
24	1142	63.5	217	9	ADY58885 Human cat

25	1138	63.3	217	7	ADDS5928 Human mat
26	1135	63.1	217	7	ADDS5930 Human mat
27	1128	62.7	217	7	ADDS5929 Human mat
28	963	53.5	330	9	ADZ11834 Canine ca
29	960.5	53.4	329	2	AA82720 Human ost
30	960.5	53.4	329	2	AA82559 Prepro-ca
31	960.5	53.4	329	2	AAW41645 Human cat
32	960.5	53.4	329	2	AAW39216 Human cat
33	960.5	53.4	329	3	AAW81000 Human pro
34	960.5	53.4	329	7	ADDE60397 Human pro
35	960.5	53.4	329	7	ADDE60397 Human pro
36	960.5	53.4	329	8	ADP12997 Protein e
37	960.5	53.4	329	8	ADQ18392 Human sof
38	960.5	53.4	329	8	ABM80141 Tumour-as
39	960.5	53.4	329	8	ADP25400 PRO polyp
40	960.5	53.4	329	9	AEC01585 Human cat
41	960.5	53.4	361	3	AA857115 Human pro
42	958.5	53.3	329	3	AA830506 A monkey
43	956.5	53.2	329	2	AAW77071 Rat Cathe
44	954	53.0	329	7	ADDE60395 Rat Prote
45	954	53.0	329	7	ADDA5859 Rat Prote

#### ALIGNMENTS

RESULT 1  
ID ADP65168 standard; protein; 331 AA.  
XX  
AC ADP65168;  
XX  
DT 12-AUG-2004 (first entry)  
XX  
DE Human cathepsin S preproprotein.  
XX  
KW autoimmune disease; arthritis; gene expression analysis;  
KW rheumatoid arthritis; collagen-induced; immunosuppressive; antineumatic;  
KW antirheumatic; osteopathic; antigenic; antinflammatory; dermatological;  
KW immunomodulatory; lupus; ankylosing spondylitis; fibrositis;  
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;  
KW immune; human.  
XX  
OS Homo sapiens.  
XX  
PN WO2003072827-A1.  
XX  
PD 04-SEP-2003.  
XX  
PF 31-OCT-2002; 2002MO-US035433.  
XX  
PR 31-OCT-2001; 2001US-0336220P.  
XX  
PA (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.  
XX  
PT Hirsch R, Thornton SJ;  
XX  
DR WPI; 2003-712740/67.  
XX  
DR GENBANK; NP\_004070.  
XX  
PT Diagnosing and analyzing autoimmune disease using gene expression  
PT profiles and microarray technology; useful for diagnosing and treating  
PT rheumatoid arthritis; lupus; fibrositis; osteoarthritis; fibromyalgia and  
PT gout.  
XX  
PS Disclosure; Page; 56pp; English.  
XX  
CC The invention relates to a novel method for diagnosing and analyzing  
CC autoimmune disease or arthritis. The method comprises obtaining a  
CC patient sample containing mRNA, analyzing gene expression using the mRNA  
CC that results in a gene expression signature of the mRNA, and using that  
CC gene expression signature to diagnose or analyse the autoimmune disease  
CC or arthritis in the patient, where gene expression of at least 60% of

1	1747	97.1	331	2	05T510	HUMAN
2	1744	96.9	331	1	06T835	HUMAN
3	1744	96.9	331	1	06T835	HUMAN
4	1701.5	94.6	330	1	CATS_SAIB	
5	1600	88.9	331	1	CATS_CANPA	
6	1406.5	78.2	341	2	06PCU5	RAT
7	1378.5	76.6	342	2	08BS25	MOUSE
8	1367.5	76.0	340	1	CATS_MOUSE	
9	1366.5	76.0	340	2	09NM14	MOUSE
10	1309.5	72.8	330	1	CATS_FAT	
11	1230.5	68.4	328	2	05ZMR6	CHICK
12	1140	63.4	333	2	06DJC1	XENTRA
13	1136.5	63.2	320	2	06DE57	XENTRA
14	1112.5	61.8	333	2	07T054	XENIA
15	1094	60.8	334	2	07T183	9C1CH
16	1070	59.5	330	2	0502A6	BRARE
17	1065	59.2	330	2	0502H6	BRARE
18	1050	58.4	330	2	0566T8	BRARE
19	1047	58.2	330	2	04ORH8	BRARE
20	1038.5	57.7	337	2	05BHF5	PAROL
21	1036	57.6	317	1	CATS_BOVIN	
22	1026	57.0	337	2	06JZ24	FUNHE
23	1022	56.8	331	2	090324	CYPCA
24	1004	55.8	334	2	04SM28	TETNG
25	983.5	55.7	327	2	06F6A0	ORYLIA
26	968.5	53.8	330	1	CATK_PIG	
27	968.5	53.8	334	2	05E9B6	BOVIN
28	965.5	53.7	329	1	CATK_RABIT	
29	961.5	53.4	329	1	CATK_MACPA	
30	961.5	53.4	329	1	CATK_MACMU	
31	960.5	53.4	329	1	CATK_HUMAN	

45	912	50.7	338	2	Q50WC2_ONCMY	Q66121 xenopus laevis
44	912.5	50.8	336	2	Q68F41_XENLA	Q68421 tetraodon n
43	914	50.7	335	2	Q68F41_XENLA	Q68421 xenopus laevis
42	919	51.1	337	2	Q4SMW2_TETNG	Q66121 xenopus laevis
41	926	51.5	336	2	Q68F61_ORYLA	Q66121 xenopus laevis
40	927.5	51.6	337	2	Q75S28_CYPCA	Q75628 cyprinus carpio
39	934	51.9	329	2	Q6XR7F_FUNDH	Q66121 xenopus laevis
38	934	51.9	329	2	Q6XR7F_FUNDH	Q66121 xenopus laevis
37	935	52.0	335	2	Q54510_MOUSE	Q54510 mus musculus
36	937	52.1	329	2	Q54510_MOUSE	Q54510 mus musculus
35	937	52.1	329	2	Q54510_MOUSE	Q54510 mus musculus
34	954	53.0	329	1	CATK_MOUSE	Q55097 xenopus laevis
33	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
32	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
31	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
30	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
29	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
28	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
27	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
26	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
25	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
24	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
23	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
22	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
21	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
20	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
19	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
18	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
17	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
16	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
15	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
14	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
13	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
12	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
11	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
10	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
9	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
8	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
7	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
6	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
5	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
4	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
3	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
2	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis
1	960.5	53.4	329	2	Q68A77_XENLA	Q68421 xenopus laevis

## ALIGNMENTS

```

RESULT 1
05T510_HUMAN
ID 05T510_HUMAN PRELIMINARY; PRT; 331 AA.
AC 05T510;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Cathepsin S.
GN Name=CTSS; ORFNames=RP11_361I22.2-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NOCLEOTIDE SEQUENCE.
RA Howden P.;
RL Submitted (MAY-2005) to the EMBL/Genbank/DBJ databases.
EMLB; AL356292; CAL13657.1; -: Genomic DNA.
DR GO; GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
SQ SEQUENCE 331 AA; 37495 MW; 86093519DB6F0269 CAC64;

```

Query Match 97.1%; Score 1747; DB 2; Length 331;

Best Local Similarity 97.0%; Pred. No. 2.5e-135;  
Matches 321; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY	I	MKQIVCVLLVCGSSAAVAOLHKDPTLDHNNHMKTYGQGXKEKKNEAVRLLIWEKNLKFVM	60
Dd	1	MKRVLVCLLVCGSSAAVAQLHKDPTLDHNNHLMKTYGQXKEKKNEAARLLIWEKNLKFVM	60
QY	61	LHNLESHMGMSHYDLGMNHLGDMTSEEVMSLSLRVPSQMQRNIITYKSNAOILLPDSVD	120
Dd	61	LHNLESHMGMSHYDLGMNHLGDMTSEEVMSLSLSLRVPSQMQRNIITYKSNAORILLPDSVD	120
QY	121	WREKGCVTEVKYQGSGCAGMAFSAVGALAEAOIKLTGKLVSLSAQNLVDCSTERYGNKC	180
Dd	121	WREKGCVTEVKYQGSGCAGMAFSAVGALAEAOIKLTGKLVSLSAQNLVDCSTERYGNKC	180
QY	181	NGCFMTRARQYIIIDNNGIDSASYYKATTDQCKQYDSRYRAATSKYTELPYGREEDYLKE	240
Dd	181	NGCFMTRARQYIIIDNKGIDSASYYKAMDQCKQYDSRYRAATCSKYTELPHYGREEDYLKE	240
QY	241	VVANLKPFVGVGDASHSPFFLYRSQVVEPCSTONVNHGVLVWGQVGLNGEKEXYLVKNSW	300
Dd	241	AANAKPFVSVGDADAHSPFFLYRSQVVEPECTQNNVNHGVLVVGIGDLNGKEXYLVKNSW	300
QY	301	GRNFGEEGYIRMARNGKHNCGIASPSPYPEI	331
Dd	301	GHNFGEEGYIRMARNGKHNCGIASPSPYPEI	331



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM proteain - nucleic search, using frame\_plus.p2n model

Run on: November 22, 2005, 11:13:59 ; Search time 806 Seconds  
(without alignments)  
3395.985 Million cell updates/sec

Title: US-10-809-816a-2  
Perfect score: 1799  
Sequence: 1 MKQLCVLLVCSSAVQLHK.....MARKNKGCIASPFPEYPI 331

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlp  
-MODEL=frame.p2n.model  
-O=/cgn2\_1/USPTO\_spool\_p/US10809816/runac\_21112005\_094203\_8378/app\_query.fasta\_1.519  
-DB=Published\_Applications\_NA\_Main -QFMT=fastac -SUFFIX=p2n.rnpbm  
-MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1  
-MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct  
-THR MAX=100 -THR MIN=0 -ALIGN=10 -MODE=LOCAL -OUTFMT=prc -NORM=ext  
-HEASize=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10809816@cgn2\_1.1549@runac\_21112005\_094203\_8378 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published\_Applications\_NA\_Main:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1799	100.0	1798	9	US-10-809-816a-1
2	1747	97.1	1784	7	US-10-641-643-1227
3	1747	97.1	4100	9	US-10-973-858-19
4	1747	97.1	4100	9	US-10-287-436a-51
5	1744	96.9	1254	8	US-10-635-398-95
6	1744	96.9	1752	9	US-10-956-157-1177
7	1744	96.9	1752	9	US-10-287-436a-953

8	1740	96.7	1643	5	US-10-099-275-1	Sequence 1, Appl1
9	1740	96.7	1763	3	US-09-873-367C-173	Sequence 173, App
10	1740	96.7	1763	3	US-10-843-641A-173	Sequence 173, App
11	1740	96.7	1763	3	US-10-287-436a-954	Sequence 554, App
12	1738	96.6	1255	6	US-10-172-118-987	Sequence 987, App
13	1738	96.6	1255	7	US-10-342-887-987	Sequence 987, App
14	1738	96.6	1255	7	US-10-283-975A-78	Sequence 78, Appl
15	1738	96.6	1255	8	US-10-278-698-206	Sequence 206, Appl
16	1738	96.6	1255	8	US-10-278-698-722	Sequence 722, App
17	1701.5	94.6	993	5	US-10-010-577-1	Sequence 1, Appl1
18	1701.5	94.6	993	5	US-10-894-104-1	Sequence 1, Appl1
19	1680	93.4	3164	7	US-10-641-643-816	Sequence 816, App
20	1680	88.9	996	8	US-10-894-106-1	Sequence 1, Appl1
21	1580	87.8	996	3	US-09-990-064-1	Sequence 1, Appl1
22	1570	87.3	1266	5	US-10-010-580-1	Sequence 1, Appl1
23	1318.5	73.3	1330	7	US-10-191-803-109	Sequence 109, App
24	1318.5	73.3	1330	7	US-10-152-319A-1700	Sequence 1700, App
25	1318.5	73.3	1330	9	US-10-486-706-131	Sequence 131, App
26	1239	68.9	1400	9	US-10-956-157-6412	Sequence 6412, App
27	1144	63.6	667	8	US-10-635-398-97	Sequence 97, Appl
28	963	53.5	1001	9	US-10-929-919A-31	Sequence 31, Appl
29	960.5	53.4	1482	6	US-10-318-584-1	Sequence 1, Appl1
30	960.5	53.4	1482	6	US-10-723-860-1210	Sequence 1210, Ap
31	960.5	53.4	1482	8	US-10-756-149-1175	Sequence 1175, Ap
32	960.5	53.4	1619	3	US-09-953-956-1	Sequence 1, Appl1
33	960.5	53.4	1619	5	US-10-114-464-1	Sequence 1, Appl1
34	960.5	53.4	1619	8	US-10-726-645-1	Sequence 1, Appl1
35	960.5	53.4	1669	6	US-10-172-118-479	Sequence 479, App
36	960.5	53.4	1669	7	US-10-240-425-1556	Sequence 1556, Ap
37	960.5	53.4	1669	7	US-10-342-887-479	Sequence 479, App
38	960.5	53.4	1669	7	US-10-956-157-408	Sequence 408, App
39	960.5	53.4	1671	5	US-10-044-090-462	Sequence 462, App
40	960.5	53.4	1702	8	US-10-278-698-204	Sequence 204, App
41	960.5	53.4	1702	8	US-10-278-698-720	Sequence 720, App
42	960.5	53.4	1702	9	US-10-973-858-17	Sequence 17, Appl
43	960.5	53.4	1756	3	US-09-925-300-753	Sequence 753, App
44	960.5	53.4	2017	5	US-10-198-846-10186	Sequence 10186, A
45	960.5	53.4	2042	8	US-10-357-930-22382	Sequence 22382, A

#### ALIGNMENTS

RESULT 1  
US-10-809-816A-1  
; Sequence 1, Application US/10809816A.  
; Publication No. US20050214774A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Shyr-Jiann et al.  
; TITLE OF INVENTION: ISOLATED MONKEY CATHEPSIN S PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING MONKEY CATHEPSIN S PROTEINS,  
; FILE REFERENCE: CLO01507  
; CURRENT APPLICATION NUMBER: US/10/809,816A  
; CURRENT FILING DATE: 2004-03-26  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 1  
; LENGTH: 1798  
; TYPE: DNA  
; ORGANISM: Cynomolgus monkey  
; US-10-809-816A-1

#### Alignment Scores:

Pred. No.: 2,176-216  
Score: 1799.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%

Length: 1798  
Matches: 331  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-10-809-816A-2 (1-331) x US-10-809-816A-1 (1-1798)  
Qy 1 MetLysGlnLeuValCysValLeuValCysSerSerAlaValAlaGlnLeuHisLys 20

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

CM protein - nucleic search, using frame\_plus.p2n model

Run on: November 22, 2005, 12:16:10 ; Search time 224 Seconds  
(without alignments)  
194.229 Million cell updates/sec

Title: US-10-809-816A-2

Perfect score: 1799

Sequence: 1 MKOLVCLIVCSSAVQLHK.....MARRKGNHGGLSPSYPEI 331

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3196817 seqs, 65720914 residues

Total number of hits satisfying chosen parameters: 6393634

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=rlp  
-Q=/cgn2\_1/USFTO.spool.p/US10809816/runat\_21112005\_094204\_8411/app.query.fasta.1.519  
-DB=Published Applications\_NA\_New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=10 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10809816 @cgn\_1\_1\_183 @runat\_21112005\_094204\_8411  
-NCPUL=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA\_New:

1: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq3:\*  
10: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	859	47.7	1547	1	US-10-131-826A-11
2	254.5	14.1	2197	1	US-10-131-826A-313
3	81	4.5	3721	1	US-10-131-826A-543
4	79.5	4.4	2760	9	US-11-008-331-1
5	78.5	4.4	4040	1	US-10-131-826A-425
6	75.5	4.2	4338	1	US-10-526-731-2
7	75.5	4.2	9719	9	US-11-042-988-10
8	74.5	4.1	2745	1	US-10-647-956A-5

9	70.5	3.9	1777	1	US-10-632-150-51
10	70.5	3.9	2602	1	US-10-510-386-229
11	69.5	3.9	1089	1	US-10-660-499A-1
12	68.5	3.8	1299	1	US-10-689-742-143
13	68.5	3.8	1799	7	US-11-071-651-13
14	68.5	3.8	2584	1	US-10-510-386-101
15	68.5	3.8	8931	7	US-11-071-651-16
16	67.5	3.8	2024	1	US-10-131-826A-259
17	67.5	3.8	3479	1	US-10-131-826A-123
18	67	3.7	1677	1	US-10-510-386-197
19	67	3.7	2237	1	US-10-510-386-445
20	67	3.7	340000	7	US-11-102-978-3
21	66.5	3.7	3048	1	US-10-957-659-50
22	66.5	3.7	3625	1	US-10-510-386-27
23	66.5	3.7	4041	1	US-10-510-386-199
24	66	3.7	1368	7	US-11-074-176-49
25	66	3.7	1670	1	US-10-614-599-2
26	65.5	3.6	993	9	US-11-082-389-415
27	65.5	3.6	1209	9	US-11-082-389-417
28	65	3.6	645	7	US-11-074-176-359
29	65	3.6	675	7	US-11-074-176-233
30	65	3.6	2347	1	US-10-510-386-23
31	64.5	3.6	2859	1	US-10-689-742-99
32	64	3.6	693	1	US-10-131-826A-499
33	64	3.6	4158	1	US-10-971-560-3
34	64	3.6	6458	1	US-10-971-560-1
35	63.5	3.5	1675	1	US-10-131-826A-135
36	63.5	3.5	3283	1	US-10-131-826A-357
37	63	3.5	2388	9	US-11-082-389-99
38	63	3.5	2388	9	US-11-082-389-101
39	63	3.5	3639	7	US-11-074-176-255
40	62.5	3.5	1095	9	US-11-082-389-439
41	62.5	3.5	4374	1	US-10-131-826A-125
42	62.5	3.5	6159	7	US-11-013-759-8
43	62.5	3.5	13943	9	US-11-012-762-39
44	62	3.4	1510	1	US-10-131-826A-399
45	62	3.4	2030	1	US-10-652-893-3

#### ALIGNMENTS

RESULT 1  
US-10-131-826A-11  
Sequence 11, Application US/10131826A  
Publication No. US20050245730A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Ausrin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C128  
CURRENT APPLICATION NUMBER: US/10/131, 826A  
CURRENT FILING DATE: 2002-04-24  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: November 22, 2005, 09:48:49 ; Search time 205 seconds  
(without alignments)  
2870.112 Million cell updates/sec

Title: US-10-809-816A-2  
Perfect score: 1799  
Sequence: 1 MKQLVCVLVCSSAQAQLMK.....MARNKGNHGHASFPSPYPER 331

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODE=frame.p2n.model -DEV=x1p  
-Q/cgn2\_1/USPTO.spool.p/US10809816/runat\_21112005\_094202\_8310/app\_query.fasta\_1.519  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=10  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10809816 @CCN\_1\_1\_290 @runat\_21112005\_094202\_8310 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Issued Patents NA.\*  
1: /cgn2\_6/prodata/1/ina/1.COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5.COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/6C.COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/PCTUS.COMB.seq.\*  
7: /cgn2\_6/prodata/1/ina/PP.COMB.seq.\*  
8: /cgn2\_6/prodata/1/ina/RP.COMB.seq.\*  
9: /cgn2\_6/prodata/1/ina/backfile1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1747	97.1	1784	3	US-09-023-655-1227
2	1740	96.7	1643	3	US-09-701-685-1
3	1701.5	94.6	993	3	US-10-010-577-1
4	1680	93.4	3164	3	US-09-023-655-816
5	1570	87.3	1126	3	US-10-010-580-1
6	960.5	53.4	1482	2	US-08-330-121B-1
7	960.5	53.4	1482	3	US-08-536-861-1
8	960.5	53.4	1482	6	PCT-US95-13820-1
9	960.5	53.4	1619	2	US-08-208-007A-1

10	960.5	53.4	1619	3	US-08-915-095A-1	Sequence 1, Appli
11	960.5	53.4	1619	3	US-08-798-096A-1	Sequence 1, Appli
12	960.5	53.4	1619	3	US-08-095A-1	Sequence 1, Appli
13	960.5	53.4	1619	3	US-09-953-956-1	Sequence 1, Appli
14	960.5	53.4	1619	3	US-08-553-125A-1	Sequence 1, Appli
15	960.5	53.4	1619	3	US-10-114-464-1	Sequence 1, Appli
16	960.5	53.4	1619	6	PCT-US94-04781-1	Sequence 1, Appli
17	960.5	53.4	1669	3	US-08-964-308-3	Sequence 3, Appli
18	960.5	53.4	1669	3	US-08-964-313-3	Sequence 3, Appli
19	960.5	53.4	1669	3	US-09-069-138-3	Sequence 3, Appli
20	956.5	53.2	990	2	US-08-806-959-1	Sequence 1, Appli
21	953.5	53.0	1614	3	US-08-684-932A-35	Sequence 35, Appl
22	951.5	52.9	990	3	US-08-964-308-13	Sequence 13, Appl
23	951.5	52.9	990	3	US-08-964-313-13	Sequence 13, Appl
24	951.5	52.9	990	3	US-09-069-138-13	Sequence 13, Appl
25	950.5	52.8	990	3	US-08-964-308-12	Sequence 12, Appl
26	950.5	52.8	990	3	US-08-964-313-12	Sequence 12, Appl
27	950.5	52.8	990	3	US-09-069-138-12	Sequence 12, Appl
28	862	47.9	1374	3	US-09-784-642-3	Sequence 3, Appli
29	859	47.7	1342	3	US-09-290-586A-2	Sequence 2, Appli
30	859	47.7	1342	3	US-09-290-586A-15	Sequence 15, Appl
31	859	47.7	1342	3	US-09-949-016-2733	Sequence 2733, Ap
32	848.5	47.2	1575	3	US-09-784-642-1	Sequence 1, Appli
33	848.5	47.2	1577	3	US-09-949-016-3909	Sequence 3909, Ap
34	848	47.1	1366	3	US-08-883-526-2	Sequence 2, Appli
35	848	47.1	1366	3	US-09-519-283-2	Sequence 2, Appli
36	803.5	44.7	1187	3	US-10-363-937-18	Sequence 18, Appl
37	796.5	44.3	1362	3	US-09-799-451-380	Sequence 380, App
38	756	42.0	1226	3	US-09-799-451-379	Sequence 379, App
39	726	40.4	1360	3	US-09-856-599-2	Sequence 2, Appli
40	664.5	36.9	1600	3	US-09-325-932A-98	Sequence 98, Appl
41	650	36.1	1056	8	US-08-813-591-1	Sequence 1, Appli
42	650	36.1	1056	8	US-09-500-651-1	Sequence 1, Appli
43	641.5	35.7	651	2	US-08-546-712-1	Sequence 1, Appli
44	641.5	35.7	651	2	US-08-751-105-1	Sequence 1, Appli
45	636	35.4	1203	3	US-09-005-298-16	Sequence 16, Appl

## ALIGNMENTS

RESULT 1  
US-09-023-655-1227  
; Sequence 1227, Application US/09023655  
; Patent No. 6607879  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; APPLICANT: Susan G. Stuart  
; APPLICANT: Jeffrey J. Selhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/023,655  
; FILING DATE: HERewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: November 22, 2005, 09:43:40 ; Search time 3721 Seconds  
(without alignments)

4161.925 Million cell updates/sec

Title: US-10-809-816A-2

Perfect score: 1799

Sequence: 1 MQOLYCVLLVCSSAVAYQLHK.....MARKKNHGCIASFSPYPEI 331

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+pn.model -DEV=xlp  
-O=/cn2.1/USPTO.spool.p/US10809816/runat.21112005.094202.8297/app.query.fasta\_1.519  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bites -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdt -LIST=45  
-LOCALIGN=200 -THR\_SCORE=PCT -THR\_MAX=100 -THR\_MIN=0 -ALIGN=10 -MODE=LOCAL  
-OUTFMT=PCT -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10809816 @CN2\_1\_1.8010 @runat.21112005.094202.8297 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEOUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: +  
1: gb\_esc1: +  
2: gb\_esc2: +  
3: gb\_esc3: +  
4: gb\_hic: +  
5: gb\_esc4: +  
6: gb\_esc5: +  
7: gb\_esc6: +  
8: gb\_esc7: +  
9: gb\_gsa1: +  
10: gb\_gsa2: +  
11: gb\_gsa3: +

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1747	97.1	1636	4	CR626534 full-length
2	1747	97.1	1636	4	CR626534 full-length
3	1747	97.1	1717	4	CR613968 full-length
4	1747	97.1	1733	4	CR612707 full-length
5	1552	86.3	896	1	AL548051 AL548051
6	1537	84.9	994	5	BX417536 BX417536
7	1519	84.4	953	1	AL513990 AL513990

8	1500.5	83.4	956	1	AL551671	AL551671
9	1486.5	82.6	1028	1	AL544870	AL544870
10	1406.5	78.2	1091	6	CF110629	Shuizom1
11	1392	77.4	986	7	CO579243	ILLUMIN
12	1386	77.0	1022	4	CR607339	full-length
13	1378.5	76.6	1355	4	AK028366	mus muscu
14	1345.5	74.8	944	8	DR772389	ILLUMIN
15	1345.5	74.8	1287	4	BC011104	mus muscu
16	1333	74.1	928	8	DN824626	ILLUMIN
17	1319.5	73.3	983	7	CO646849	ILLUMIN
18	1319	73.3	945	7	CO646666	ILLUMIN
19	1316	73.2	976	5	BX345726	BX345726
20	1304	72.5	848	6	BG339651	BG339651
21	1292	71.8	901	5	BW972127	BW972127
22	1292	71.8	901	5	DN824129	LEB0011.C2
23	1251	69.5	934	5	BO710837	AGENCOURT
24	1247	69.3	908	5	BW975020	BW975020
25	1245	69.2	879	8	DN824771	LEB0013.C2
26	1238	68.8	876	8	DN824658	LEB0013.C2
27	1235	68.6	1068	3	BM806785	AGENCOURT
28	1234	68.6	795	2	BG541942	602569445
29	1231	68.4	839	1	AU140074	AU140074
30	1218	67.7	942	3	BP156785	BP156785
31	1215	67.5	789	1	AM035440	AM035440
32	1213	67.4	864	8	DN823937	LEB0011.C2
33	1209	67.2	764	8	DN97412	TC117580
34	1209	67.2	855	8	DN824864	LEB0014.C2
35	1185	65.9	758	3	BT766075	603053051
36	1181	65.6	898	7	CJ011620	CJ011620
37	1179	65.5	759	7	CK973194	4104166.B
38	1175	65.3	782	2	BG548684	602576491
39	1168	64.9	748	7	CK944131	4068178.B
40	1161	64.5	809	6	CB993383	AGENCOURT
41	1160	64.5	732	6	CB986537	AGENCOURT
42	1150	63.9	710	6	CB986513	AGENCOURT
43	1150	63.9	853	2	BE531143	601278324
44	1145	63.6	702	8	DN272753	DN272753
45	1144	63.6	1048	2	BE616367	601279178

#### ALIGNMENTS

RESULT 1  
LOCUS CR626534 1636 bp mRNA linear HTC 21-JUL-2004  
DEFINITION full-length cDNA clone CSOD1034YM16 of Piacenta Cot 25-normalized.  
ACCESSION CR626534.1 GI:50507341  
VERSION CR626534.1  
KEYWORDS HTC; CNSLT\_cDNA.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
1 (bases 1 to 1636)  
Li, W.B., Gruber, C., Dessee, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Feng Liang Email: fliang@life.technet.com URL:  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paraday Avenue  
2 (bases 1 to 1636)  
Genoscope.  
Direct Submission  
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage:  
BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr  
- Web: www.genoscope.cns.fr)  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: November 22, 2005, 08:23:59 ; Search time 573 Seconds

(without alignments)  
3849,936 Million cell updates/sec

Title: US-10-809-816A-2

Sequence: 1 MKQLVCLVLCSSAVNAQLHK.....MARKNKHGIAFPSPYPP 331

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 333346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=slp  
-G=gen2.1/USPTO.spool\_p/US10809816/runat\_21112005\_094201\_8278/app\_query.fasta.1.519  
-DB=N.GeneSeq -QFMT=faa1ap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNIT5=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=10 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10809816 @CCN 1.1 1096 @runat\_21112005\_094201\_8278 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEBUFFER -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLCG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.GeneSeq\_21.\*

1: geneeqn19806:\*  
2: geneeqn19908:\*  
3: geneeqn20008:\*  
4: geneeqn20018:\*  
5: geneeqn20028:\*  
6: geneeqn20038:\*  
7: geneeqn20048:\*  
8: geneeqn20058:\*  
9: geneeqn20068:\*  
10: geneeqn20078:\*  
11: geneeqn20088:\*  
12: geneeqn20098:\*  
13: geneeqn20108:\*  
14: geneeqn20118:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1747	97.1	1784	AD131901 Human CDN
2	1747	97.1	1784	AD883968 Human lym
3	1747	97.1	4100	ADP64970 Human cat
4	1747	97.1	4100	ADY14945 DNA encod

5	1747	97.1	4100	ADY14943	ADY14943 DNA encod
6	1747	97.1	4100	ADY14943	ADY14943 DNA encod
7	1747	97.1	4100	ADY14943	ADY14943 DNA encod
8	1747	97.1	4100	ADY14943	ADY14943 DNA encod
9	1747	97.1	4100	ADY14943	ADY14943 DNA encod
10	1747	97.1	4100	ADY14943	ADY14943 DNA encod
11	1747	97.1	4100	ADY14943	ADY14943 DNA encod
12	1747	97.1	4100	ADY14943	ADY14943 DNA encod
13	1747	97.1	4100	ADY14943	ADY14943 DNA encod
14	1747	97.1	4100	ADY14943	ADY14943 DNA encod
15	1747	97.1	4100	ADY14943	ADY14943 DNA encod
16	1747	97.1	4100	ADY14943	ADY14943 DNA encod
17	1747	97.1	4100	ADY14943	ADY14943 DNA encod
18	1747	97.1	4100	ADY14943	ADY14943 DNA encod
19	1747	97.1	4100	ADY14943	ADY14943 DNA encod
20	1747	97.1	4100	ADY14943	ADY14943 DNA encod
21	1747	97.1	4100	ADY14943	ADY14943 DNA encod
22	1747	97.1	4100	ADY14943	ADY14943 DNA encod
23	1747	97.1	4100	ADY14943	ADY14943 DNA encod
24	1747	97.1	4100	ADY14943	ADY14943 DNA encod
25	1747	97.1	4100	ADY14943	ADY14943 DNA encod
26	1747	97.1	4100	ADY14943	ADY14943 DNA encod
27	1747	97.1	4100	ADY14943	ADY14943 DNA encod
28	1747	97.1	4100	ADY14943	ADY14943 DNA encod
29	1747	97.1	4100	ADY14943	ADY14943 DNA encod
30	1747	97.1	4100	ADY14943	ADY14943 DNA encod
31	1747	97.1	4100	ADY14943	ADY14943 DNA encod
32	1747	97.1	4100	ADY14943	ADY14943 DNA encod
33	1747	97.1	4100	ADY14943	ADY14943 DNA encod
34	1747	97.1	4100	ADY14943	ADY14943 DNA encod
35	1747	97.1	4100	ADY14943	ADY14943 DNA encod
36	1747	97.1	4100	ADY14943	ADY14943 DNA encod
37	1747	97.1	4100	ADY14943	ADY14943 DNA encod
38	1747	97.1	4100	ADY14943	ADY14943 DNA encod
39	1747	97.1	4100	ADY14943	ADY14943 DNA encod
40	1747	97.1	4100	ADY14943	ADY14943 DNA encod
41	1747	97.1	4100	ADY14943	ADY14943 DNA encod
42	1747	97.1	4100	ADY14943	ADY14943 DNA encod
43	1747	97.1	4100	ADY14943	ADY14943 DNA encod
44	1747	97.1	4100	ADY14943	ADY14943 DNA encod
45	1747	97.1	4100	ADY14943	ADY14943 DNA encod

#### ALIGNMENTS

RESULT 1	AD131901	AD131901 standard; cDNA; 1784 BP.
ID	AD131901	AD131901 standard; cDNA; 1784 BP.
XX	AD131901	AD131901 standard; cDNA; 1784 BP.
AC	AD131901	AD131901 standard; cDNA; 1784 BP.
XX	AD131901	AD131901 standard; cDNA; 1784 BP.
DT	17-JUN-2004	(first entry)
XX	17-JUN-2004	(first entry)
DE	Human CDNA #1227.	
XX	Human CDNA #1227.	
KW	Human; gene; ss; immunological response; immunopathological condition;	
KW	Crown's disease; asthma; ulcerative colitis; hyperostosis; arthritis;	
KW	irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;	
KW	acute monocytic leukemia; antiinflammatory; antiaesthetic; antitumor;	
KW	osteopathic; antiarthritic; antirheumatic; cytostatic.	
OS	Homo sapiens.	
XX	Homo sapiens.	
PN	US6607879-B1.	
XX	US6607879-B1.	
PD	19-AUG-2003.	
XX	19-AUG-2003.	
PF	09-FEB-1998;	98US-00023655.
XX	09-FEB-1998;	98US-00023655.
PR	09-FEB-1998;	98US-00023655.
XX	09-FEB-1998;	98US-00023655.
PA	(INCY-) INCYTE CORP.	
XX	(INCY-) INCYTE CORP.	



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: November 22, 2005, 09:26:44 ; Search time 4594 Seconds  
(without alignments)  
4095.599 Million cell updates/sec

Title: US-10-809-816a-2  
Perfect score: 1799  
Sequence: 1 MKQLVCVLIVCSSAVQAHLK.....MARKNKGHCIAISFSPYPEI 331

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus.p2n.model -DEV=xlp  
-G/cgn2.1/USFTO.spool.p/US10809816/runcat\_21112005\_094201\_8286/app\_query.fasta.1.519  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -DOFCL=0 -DOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=10 -MODE=LOCAL  
-OUTFMT=pcr -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10809816 @CCN\_1\_1\_7415@runcat\_21112005\_094201\_8286 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -KONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:.\*  
1: gb\_da:.\*  
2: gb\_in:.\*  
3: gb\_env:.\*  
4: gb\_om:.\*  
5: gb\_ov:.\*  
6: gb\_par:.\*  
7: gb\_ph:.\*  
8: gb\_pr:.\*  
9: gb\_ro:.\*  
10: gb\_scs:.\*  
11: gb\_sy:.\*  
12: gb\_un:.\*  
13: gb\_vl:.\*  
14: gb\_hcg:.\*  
15: gb\_pl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Match	Length	DB	ID	Description
1	1747	97.1	996	11	AY888599	Synthetic
2	1747	97.1	996	11	AY891245	Synthetic
3	1747	97.1	996	11	AY893335	Synthetic

Result No.	Score	Match	Length	DB	ID	Description
4	1747	97.1	1259	6	C0728273	Sequence
5	1747	97.1	1784	6	AR380682	Sequence
6	1747	97.1	4100	6	CS031243	Sequence
7	1747	97.1	4100	6	CS031245	Sequence
8	1747	97.1	4100	6	CS040195	Sequence
9	1747	97.1	4100	6	CS040197	Sequence
10	1747	97.1	4100	6	CS118410	Sequence
11	1744	96.9	996	8	CR541676	Sequence
12	1744	96.9	996	11	AY893542	Synthetic
13	1744	96.9	1743	8	BC002642	Sequence
14	1740	96.7	1643	6	BD245063	Sequence
15	1740	96.7	1643	6	AR209918	Sequence
16	1740	96.7	1763	6	AX329664	Sequence
17	1740	96.7	1763	8	HDMCATS	Sequence
18	1738	96.6	1255	6	CS036083	Sequence
19	1738	96.6	1255	6	CS045035	Sequence
20	1738	96.6	1255	6	AX774762	Sequence
21	1738	96.6	1255	8	HDMCATS	Sequence
22	1738	96.6	1255	8	S93414	Sequence
23	1701.5	94.6	993	6	AR579499	Sequence
24	1701.5	94.6	993	6	AX774253	Sequence
25	1701.5	94.6	1143	8	AY156691	Sequence
26	1680	93.4	3164	6	AR380271	Sequence
27	1600	88.9	996	6	AX797393	Sequence
28	1600	88.9	1125	4	AY156692	Sequence
29	1570	87.3	1126	6	AR568030	Sequence
30	1406.5	78.2	1191	9	BC059142	Sequence
31	1366.5	76.0	1279	9	BC002125	Sequence
32	1361.5	75.7	1305	9	AF038546	Sequence
33	1349.5	75.0	1296	9	MCATHEPS	Sequence
34	1318.5	73.3	1330	9	RATCTHS	Sequence
35	1230.5	68.4	1499	5	AJ719318	Sequence
36	1140.5	63.4	1752	5	BC077285	Sequence
37	1140	63.4	1778	5	BC075261	Sequence
38	1112.5	61.8	2248	5	BC056059	Sequence
39	1094	60.8	1088	5	AY333300	Sequence
40	1070	59.5	1242	5	BC095788	Sequence
41	1065	59.2	1246	5	BC095694	Sequence
42	1053	58.5	601	10	BY179264	Sequence
43	1050	58.4	1338	5	BC093339	Sequence
44	1047	58.2	1257	5	BC096862	Sequence
45	1043	58.0	1399	5	AY622858	Sequence

## ALIGNMENTS

RESULT 1  
AY888599  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AY888599 996 bp mRNA linear SYN 22-MAR-2005  
Synthetic construct Homo sapiens clone FHM019213.01x cathepsin S (CTSS) mRNA, complete cds.  
AY888599  
AY888599.1 GI:61358281  
Human ORF Project.  
synthetic construct  
synthetic construct  
Other sequences; artificial sequences.  
1 (bases 1 to 996)  
Hines, L., Rolfs, A., Jepsen, D., Moreira, D., Raphael, J., Kelley, F., Shen, B., Halleck, A., Koundinya, M., Hu, Y., Zuo, D., Tayher, E., Williamson, J. and Labaer, J.  
Cloning of human full-length CDS in Creator (TM) recombinational vector system  
Unpublished  
2 (bases 1 to 996)  
Hines, L., Rolfs, A., Jepsen, D., Moreira, D., Raphael, J., Kelley, F., Shen, B., Halleck, A., Koundinya, M., Hu, Y., Zuo, D., Tayher, E., Williamson, J. and Labaer, J.  
Direct Submission  
Submitted (04-JAN-2005) Biological Chemistry and Molecular Pharmacology, Harvard Institute of Proteomics, 320 Charles St., Cambridge, MA 02141, USA  
This CDS clone is a part of a collection of human full-length



1 MKQLVCVLVCSSAVAQLHK